知识管理论坛　ISSN 2095-5472　CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

2000–2018年国内机构知识库研究可视化分析

◎陈和　林静
厦门大学图书馆　厦门 361005

摘要：［目的 / 意义］目前IR发展到相对成熟阶段，对国内IR发展情况进行梳理、分析和总结，可为下一阶段IR发展提供参考建议。［方法 / 过程］利用 Excel、VOSviewer、CiteSpace 等软件和编程处理数据方法，对 2000 – 2018 年 CNKI 收录的 1 362 篇 IR 期刊论文进行可视化分析。［结果 / 结论］国内IR发文经历了迅速增长期、相对平稳期以及快速下滑期，预示目前国内IR研究进入新阶段，出现新的研究点；国内研究IR的高频作者（群）以中国科学院系统群体和高校系统群体为主，前者的研究比后者更为前沿和深入；《图书情报工作》《现代情报》《图书馆学研究》等是刊载IR研究论文最多的期刊，是国内IR研究领域较有影响力的期刊；国内IR研究热点包括系统构建、资源建设、学术交流、知识管理与服务、版权保护等；IR未来的可能研究热点包括科研数据、科研服务、辅助科研决策、数据增值服务等。

关键词：机构知识库；IR；数据可视化；知识图谱；文献计量；研究热点
分类号：G251

引用格式：陈和，林静. 2000-2018年国内机构知识库研究可视化分析[J/OL]. 知识管理论坛, 2019, 4(5): 276-286[引用日期]. http://www.kmf.ac.cn/p/184/.

20世纪90年代末，机构知识库（Institutional Repository，IR）概念在国外开始萌芽。2001年，美国俄亥俄州立大学图书馆和首席信息官办公室联合推出了知识库服务（Knowledge Bank Service），员工、教师和学生可以将自己的作品上传到知识库中，并长期保存，这是实际意义上IR的早期雏形[1]。2002年，惠普公司与美国麻省理工大学合作，研发了IR的构建软件——DSpace[2]，并建立麻省理工大学IR——DSpace@MIT[3]，成为第一个正式建立IR的学术机构。2004年，浙江大学图书馆基于DSpace构建了国内第一家实验性的IR。2006年8月，厦门大学图书馆正式对外公开并发布基于DSpace构建的厦门大学学术典藏库[4]，成为我国大陆

基金项目：本文系福建省中青年教师教育科研项目“基于本地博硕士学位论文系统的数据挖掘与可视化模式研究”（项目编号：JZ170312）研究成果之一。
作者简介：陈和（ORCID: 0000-0002-4506-5612），副研究馆员，硕士；林静（ORCID:0000-0002-7886-6109），副研究馆员，博士，通讯作者，E-mail: linjingxmulib@xmu.edu.cn。
收稿日期：2019-05-08　　发表日期：2019-09-04　　本文责任编辑：刘远颖

知识管理论坛，2019(5):276–286
DOI：10.13266/j.issn.2095–5472.2019.027

地区第一个正式运行 IR 的高校。中国科学院（以下简称“中科院”）国家科学图书馆兰州分馆于 2006 年基于 DSpace 进行主题学科的 IR 平台建设，并在不断完善后，在中科院系统内推广运用 [5]。此后，国内越来越多的学者开始关注和研究 IR，IR 也成为各重要图情学术会议的交流讨论主题之一 [6]，并且从 2013 年开始，中国 IR 学术研讨会每年举办一次，交流和讨论 IR 最新研究成果，以推进 IR 在中国的发展 [7]。时至今日，IR 在国内发展将近 20 年，笔者通过对国内有关 IR 的期刊论文进行文献计量和数据可视化，对 IR 研究人员、IR 研究热点等相关主题进行梳理和分析，为研究 IR 的相关人员提供参考，为今后 IR 的创新发展提出建议。

1 数据来源与处理

笔者选用 CNKI（中国知网）的期刊论文为数据源。在期刊论文检索页面，选择检索方式为“主题”，检索词为“机构知识库”“机构典藏库”“机构库”“机构仓储”“机构存储”，检索词之间的关系为逻辑或，检索期限为 2000 年到 2018 年，其他为默认选项，获得 1 648 篇文章（检索日期为 2019 年 2 月 15 日）。通过人工剔除与本研究无关的结果，如物流仓储、dSPACE 仿真系统、IR 会议通知、征稿等，最终得到有效文章篇数为 1 362 篇。由于受到 CNKI 单次下载元数据条目的限制，以上元数据是通过多个时间段检索多次汇总而来。为方便数据分析和利用数据可视化软件，数据导出格式包括了 Refworks、EndNote 和自定义格式。

2 论文发表时间分布分析

对检索获得的 1 362 篇文献的发表年份元数据字段进行统计，利用 Excel 绘制成逐年发文数量曲线图，如图 1 所示：

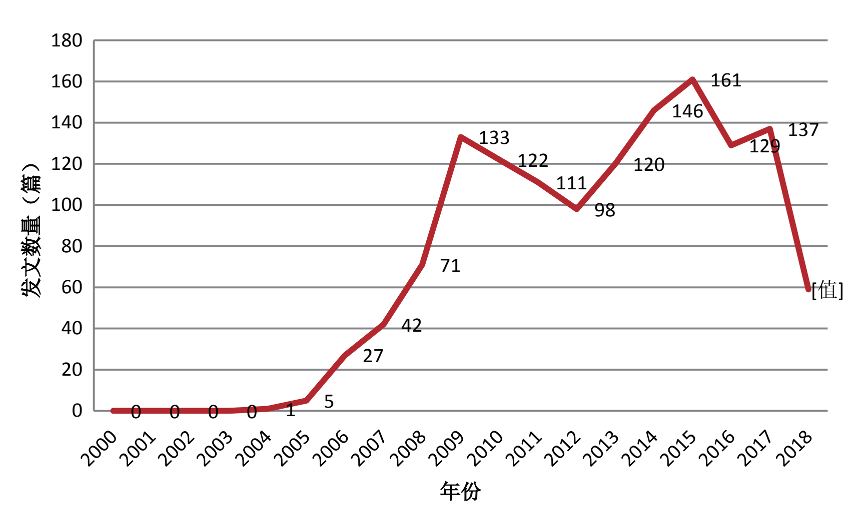


图 1 逐年发文数量曲线图

国外在2002年左右开始报导有关IR研究[8]，从图 1 可以看出，国内从 2004 年开始第一次出现 IR 研究的文章 [9]，反映了我国学者的研究紧跟国外研究步伐的特点，虽然有些时滞，但很短。从发文趋势情况来看，2004 年算是中国大陆研究 IR 的元年，从此国内 IR 进入蓬勃发展的阶段。从 2004 年到 2018 年，有关 IR 的发文情况大概可以分为三个阶段：

第一阶段，从 2004 年到 2009 年。该阶段是 IR 发文的快速增长阶段，2009 年达到第一个高峰，发文量达 133 篇。在该阶段，IR 作为新的学术课题被引入国内，引起国内图书情报



界学者极大的关注和研究热情，纷纷发表了大量的关于 IR 的学术论文。纵览这时期的论文内容，主要是介绍域外 IR 的研究进展或成果，包括 IR 的出现背景、概念内涵、学术交流功能意义等，以及 IR 作为“泊来品”对我国图书馆建设与学术研究的启示等。

第二阶段，从 2009 年到 2017 年。该阶段是 IR 发文相对稳定阶段。经过第一个阶段的学习积累，国内图书情报界对 IR 有了比较充分的理论认识，开始进入 IR 实践发展阶段，这阶段的发文涉及 IR 构建实践、IR 资源建设探索、文献计量分析、开放获取政策、学科服务、科研管理等。该阶段是 IR 在国内真正落地生根发展阶段，研究方向多样化，实现 IR 研究百花齐放式的发展，每年的发文量几乎都是百篇以上。

第三阶段，2018 年。从发文曲线来看，2018 年的总发文量比往年有比较大的下滑（原文检索时间于 2 月份，由于 CNKI 上线论文存在时滞，2019 年 6 月份本文修改时 2018 年的发文量已增至 88 篇），从趋势来看，可能还将继续下滑。该现象直观地表明，题名、关键词、摘要等字段含有“机构知识库”及相近概念的期刊论文开始减少，但是否意味着国内 IR 研究开始停滞不前或进展缓慢？笔者以 IR 当前热门话题“科研数据”为检索词在 CNKI 进行检索，其中篇名含有该词并且检索时间为 2018 年的检索结果是 78 篇，这些被检索出的期刊论文在相关字段中多数没有出现“机构知识库”及相近概念词语，但是此类论文仍然属于 IR 范畴。因此，进入 2018 年，国内 IR 研究并没有停滞，只是以“机构知识库”及相近概念作为发文主题的论文开始减少，预示国内 IR 的研究开始进入新阶段，出现新的研究热点，仅用“机构知识库”及相近概念已不能满足检索该领域的全部论文要求，关于新研究热点将在下文“研究热点分析”部分继续分析。

3 主要刊发期刊分析

经统计，1 362 篇文献刊载于 224 种期刊，依据发文量由多到少排序，并根据“二八定律”，取前 20%（45 本）期刊进行统计分析，得到如图 2 所示的比例示意图（低于 4% 的期刊作了归并处理）。

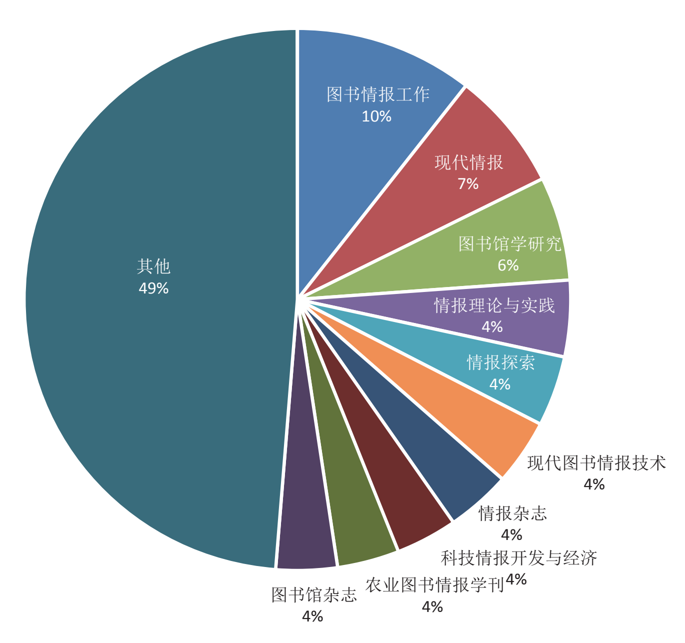


图 2 主要刊发期刊分析

知识管理论坛，2019(5):276-286
DOI：10.13266/j.issn.2095-5472.2019.027

如图 2 所示，《图书情报工作》《现代情报》《图书馆学研究》《情报理论与实践》《情报探索》《现代图书情报技术》（2017 年更名为《数据分析与知识发现》）、《情报杂志》《科技情报开发与经济》《农业图书情报学刊》《图书馆杂志》10 种期刊刊载的有关 IR 文章达到所有 IR 发文量的 50%，也就是说国内研究 IR 的作者将大部分文章投向了这些期刊。从统计结果来看，上述刊物可以视为国内 IR 研究领域的高影响力期刊，若有需要了解国内 IR 具体研究情况的用户，可以把这些刊物发表的 IR 论文作为重点关注论文。

4 论文作者分析

4.1 高频作者

对 1 362 篇论文的作者进行统计和排序，其中发文量大于等于 8 篇的作者共有 21 位，可以认为这些作者是国内 IR 研究领域的高频作者，见图 3。其中，发文频率最多的作者是祝忠明，以 34 篇高居榜首，是排在第二位作者发文量的 2 倍还多，是国内 IR 研究领域发文量超高频作者。从作者对应的工作单位来看，中科院系统的作者约占高频作者数的一半，而其他的高频作者主要来自高校系统，这从另一侧面反映了中科院系统对 IR 的研究比高校系统更加活跃，从他们的发文内容来看，中科院系统作者对 IR 的研究也比较系统、深入和前沿。

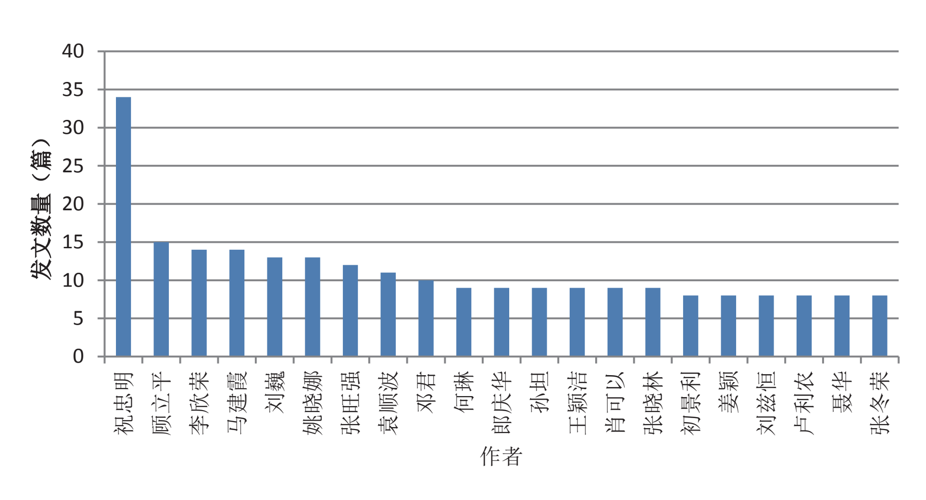


图 3 发文高频作者

4.2 作者合作网络

对原始检索结果数据的作者字段进行 Python 编程转换处理成 ris 文件，再利用可视化软件 VOSviewer 创建作者合作网络图（作者最少发文量阈值设定为 3 篇），如图 4 所示。

图 4 显示，合作成员超过 5 名的研究团队共有 4 个：①祝忠明研究团队。该团队成员基本来自中科院国家科学图书馆兰州分馆。通过他们发表的文章可以发现，该团队是中科院 IR 系统的技术开发和维护团队，他们在多个基金项目的资助下，完成了多项有关 IR 的研究。在 DSpace 作为原型系统的基础上，开发了适合中科院实际需求的新的 IR 系统——CSpace，并在全院推广使用。根据用户需求，该团队对系统持续了进行开发、更新和维护，目前已经开始将其推广到中科院系统外的国内高校系统 [10]。该团队对 IR 研究最深入、最全面、最前沿，是国内研究 IR 领域的权威团队。②顾立平研究团队。团队成员主要来自中科院国家科学图书馆。通过他们发表的文章可以发现，他们和祝忠明团队同属于中科院机构知识团队，两者相辅相成，但前者主要侧重于中



知识管理论坛
2019 年 第 5 期(总第 23 期)

科院 IR 系统的推广与服务、开放获取政策研究、辅助系统支持等。③聂华研究团队。成员主要来自北京大学图书馆。通过他们发表的文章可以发现，北京大学图书馆很早就已经介入 IR 的建设，作为 Calis-IR 项目的主要负责单位，担负了 CALIS 成员馆的 IR 建设与推广工作，现在正在推动中国高校 IR 联盟建设 [11-12]。④李国俊研究团队。团队成员主要来自北京科技大学图书馆，他们主要以北京科技大学 IR 为基础 [13]，讨论了 IR 系统构建、数字资源建设、基于 IR 的文献收录引用及计量服务研究等。

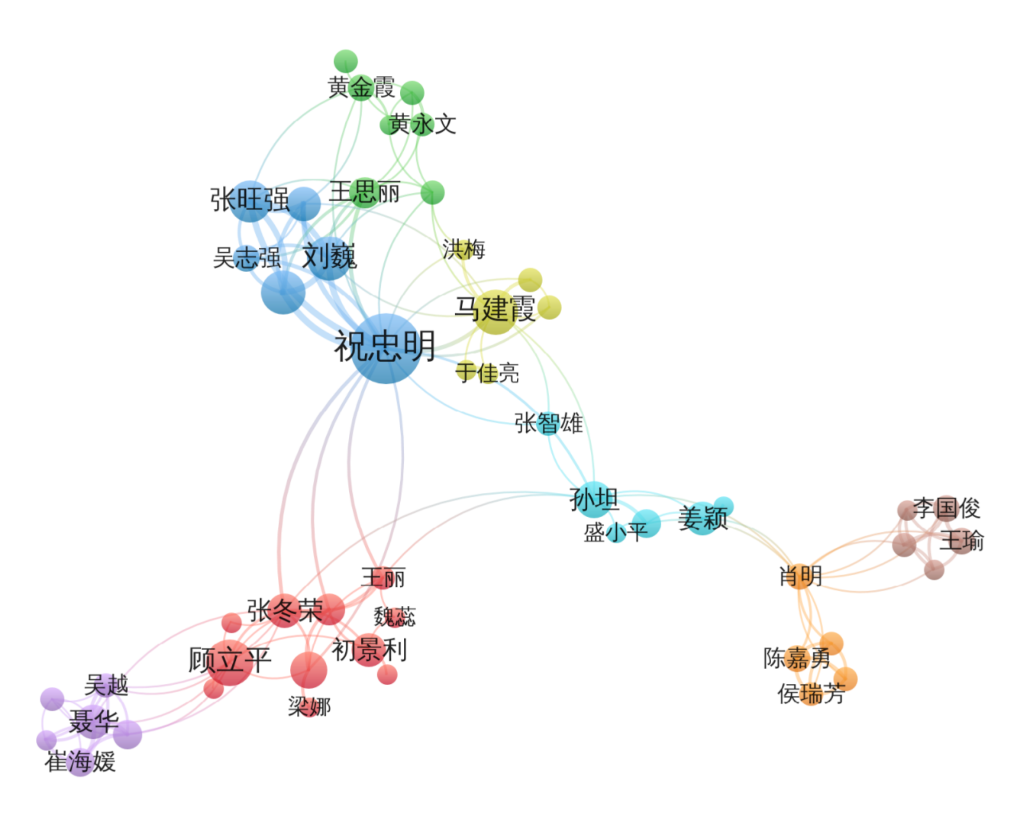


图 4 作者合作网络图

另外，从图 4 中还可以看到，孙坦、肖明在作者合作网络中起到了联结桥梁作用，联结了中科院、北京师范大学、北京邮电大学、北京科技大学等之间的校际 IR 研究合作。

5 研究热点分析

论文关键词是作者选用的最能体现论文研究内容或主题的简短词语，通过对国内有关 IR 论文的关键词进行计量分析，可以了解国内 IR 的主要研究热点。

由于关键词并不是规范的主题词，对关键词进行统计分析前，需要对关键词进行必要的归并处理，比如“开放存取”统一为“开放获取”，“机构库”“机构仓储”“机构仓储库”“机构存储”“机构存储库”“机构典藏库”“机构知识库”等统一为“机构知识库”，“大学”统一为“高校”。

5.1 高频关键词

通过对规范后的关键词进行词频统计，获得高频关键词列表，如表 1 所示（只截取前 30 个）：

知识管理论坛，2019(5):276-286
DOI：10.13266/j.issn.2095-5472.2019.027

表 1 高频关键词列表（前 30）

关键词	频次	关键词	频次	关键词	频次
机构知识库	1 057	知识服务	28	著作权	16
开放获取	367	资源建设	28	策略	15
图书馆	236	版权	23	数据管理	15
高校	214	自存储	22	学科服务	15
DSpace	59	质量控制	20	元数据	15
学术交流	38	资源共享	20	知识产权	15
数字图书馆	31	可持续发展	18	知识图谱	15
知识管理	29	美国	17	开源软件	14
IR联盟	28	内容建设	17	日本	14

如表 1 所示，由于本次研究对象是机构知识库，所以该词出现频率最高，而且远高于其他关键词。接着出现的高频关键词是“开放获取”“图书馆”“高校”，这 3 个词的出现频次也是远高于其他高频关键词。IR 是在开放获取运动迅速发展的背景下产生，是后者的一种具体实践方式。讨论开放获取时也会讨论 IR，而讨论 IR 时也往往会提及开放获取相关内容和背景。因此，“开放获取”成为 IR 之后的次超高频关键词。“图书馆”和“高校”关键词说明目前国内研究 IR 的主要是图书馆，而且以高校图书馆为主。IR 的一个重要使命是长期保存本机构的知识成果，与图书馆保存人类知识之使命相契合，可以说图书馆积极研究 IR、构建 IR、推广 IR 服务是一项职业使命，也是历史使命。构建 IR 系统的开源软件 DSpace，在推动 IR 在世界范围内广泛普及与发展方面贡献巨大，国内学者研究 IR 往往是从研究 DSpace 开始，因此，DSpace 也出现在高频关键词的前列。

对于其他的高频关键词，笔者将结合下面的关键词共现网络图进行分析。

5.2 关键词共现网络

参照作者合作网络，利用 VOSviewer 创建关键词共现网络图（见图 5）。由于“机构知识库”“开放获取”“图书馆”“高校”在该网络中属于超高频关键词，在共现网络图中覆盖了其他关键词的共现关系，因此未把它们并入共现网络图。在参数设置上，关键词出现的最小频率（mininum number of occurrences of a keywords）是 8。

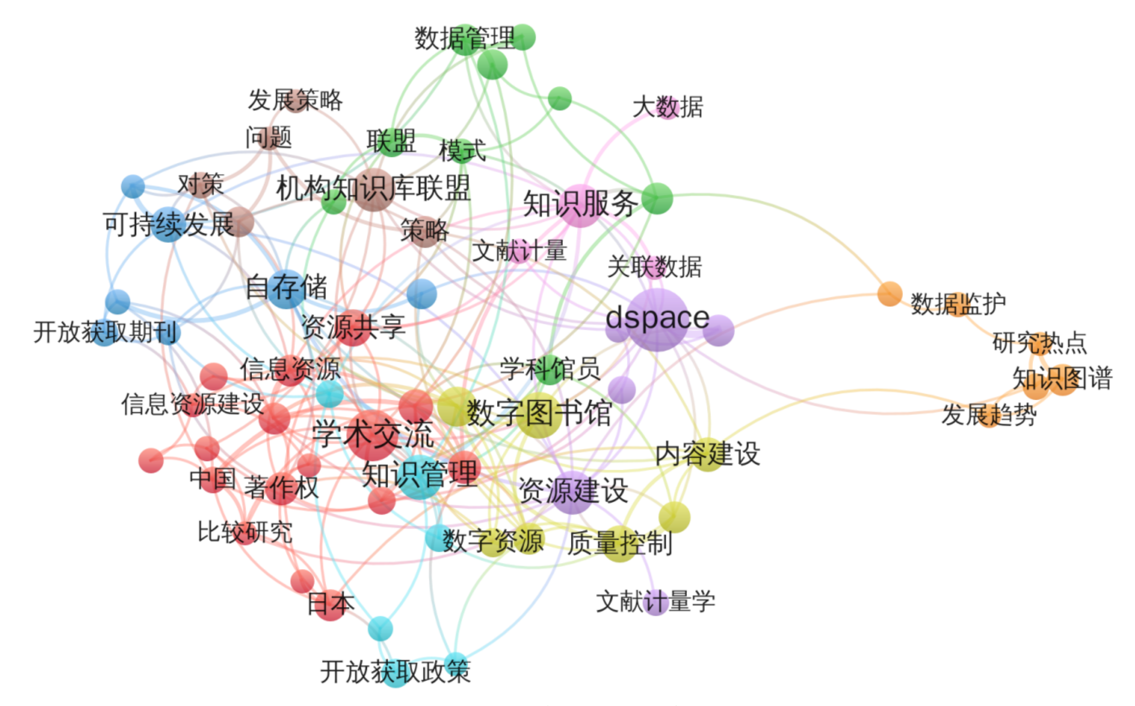


图 5 关键词共现网络



知识管理论坛
2019 年 第 5 期(总第 23 期)

通过表 1 和图 5，并结合具体文献进行综合分析，可以概括出国内 IR 领域研究的几个热点研究领域：

（1）IR 系统构建研究。该研究热点主要出现在 IR 研究早期。构建 IR 有多款软件系统，其中 DSpace 软件因其源代码开放、无偿免费使用、功能丰富、社区支持等特点吸引了无数 IR 关注者的兴趣，它实现了 IR 从理论到落地实践的过程。从发表的文献来看，内容涉及 DSpace 软件的功能架构介绍、与其他 IR 构建软件对比研究、DSpace 本地化与具体应用实践等。与之共现的主要关键词除了之前的超高频关键词外，还有数据模型、模式、元数据、互操作、关联数据、开源软件等。

（2）IR 资源建设研究。开展 IR 服务，其最根本还是要落实到具体资源的收集和管理。机构的学术成果或学术产出形式多样，包括但不限于期刊论文、会议论文、学位论文、演示文稿、专利、软件、影像视频、科学数据等。如何对这些学术成果进行有效的收集、组织和整理，是国内 IR 研究一直关注的热点问题。IR 资源建设相关的关键词有内容建设、质量控制、学术资源、开放获取资源、知识产权、自存储、服务模式、策略、文献计量学、建设模式、可持续发展、DSpace 等。

（3）IR 促进学术交流功能研究。随着信息技术与网络技术的发展以及人们对学术研究活动观念与方法的改变，传统的耗时、耗力与耗钱的学术交流模式已经不适应自由快速学术交流的趋势。开放获取的理念是让用户通过公共网络快速无障碍地获取所需要的学术资源，改变了传统的学术交流模式 [14]，IR 作为开放获取模式下的学术产物，实现了学术自由快速交流的作用，受到科研人员和普通用户极大的关注和认可。与学术交流共现的关键词有知识管理、信息资源、资源共享、IR 联盟、学术 IR、DSpace、模式、日本等。

（4）IR 知识管理与服务研究。IR 的本质是对机构的知识产出进行有效的组织和管理，并在网络上开放这些知识产品，实现学术自由交流。要开展 IR 服务，除了对原始学术资源进行收集和积累，图书馆还需要联合机构相关组织部门，如高校科研管理部门、人事管理部门等，对 IR 的内容进行有效的组织和管理，以便为不同科研人员或实体提供高效的服务 [15]。这是对 IR 更高层级的研究。与 IR 知识管理与服务相关的关键词有知识管理、知识服务、科研管理、科研成果、数据管理、大数据、资源共享、学科馆员、关联数据、IR 联盟、可持续发展等。

（5）版权保护研究。IR 是在开放获取运动的背景下产生，旨在提供学术自由交流的平台，为用户提供学术资源的无障碍访问和获取服务，但并不意味着学术资源可以不受保护地被任意使用。如何在收集机构学术成果和对外提供服务时，保护作品和作者的版权权益，同时为用户提供最大的使用便利，也是国内 IR 研究的一个研究热点。与版权保护相关的关键词有知识产权、著作权、自存储、自存档、质量控制、知识共享、合理使用、数字图书馆、对策等。

总体上，上述研究热点之间的区隔并不太明显，而是互相渗透和交融，研究热点整体上比较平均，并没有特别突出的研究热点。

5.3 关键词时间线视图

利用 CiteSpace（5.3.R4.8.31.2018）对从 CNKI 下载的以 Refworks 格式保存的元数据中的关键词节点进行时间线视图（Timeline View）分析，得到图 6。同样，为避免超高频关键词对其他关键词造成覆盖，“机构知识库”“开放获取”“图书馆”“高校”等未并入分析。

如图 6 所示，CiteSpace 根据自带算法，把关键词自动归为 9 个类别，其中 #0 机构知识库类的时间轴线（实线部分）起源最早，较完整地展示了 IR 在国内的发展过程，从 2004 年介绍引入构建 IR 系统的开源软件 DSpace 开始，到对 DSpace 进行本地化处理，以适应国内的应用环境，再到学习参考较早介入 IR 建设而且又与大陆同文同种的台湾地区高校 IR 建设经验。随

知识管理论坛，2019(5):276-286
DOI：10.13266/j.issn.2095-5472.2019.027

着 IR 的推广普及，IR 建设策略、建设模式、服务模式等主题相继被学者们探讨。在这个过程中，域外比较先进的 IR 建设实践（如日本、美国等）继续被介绍和引入。

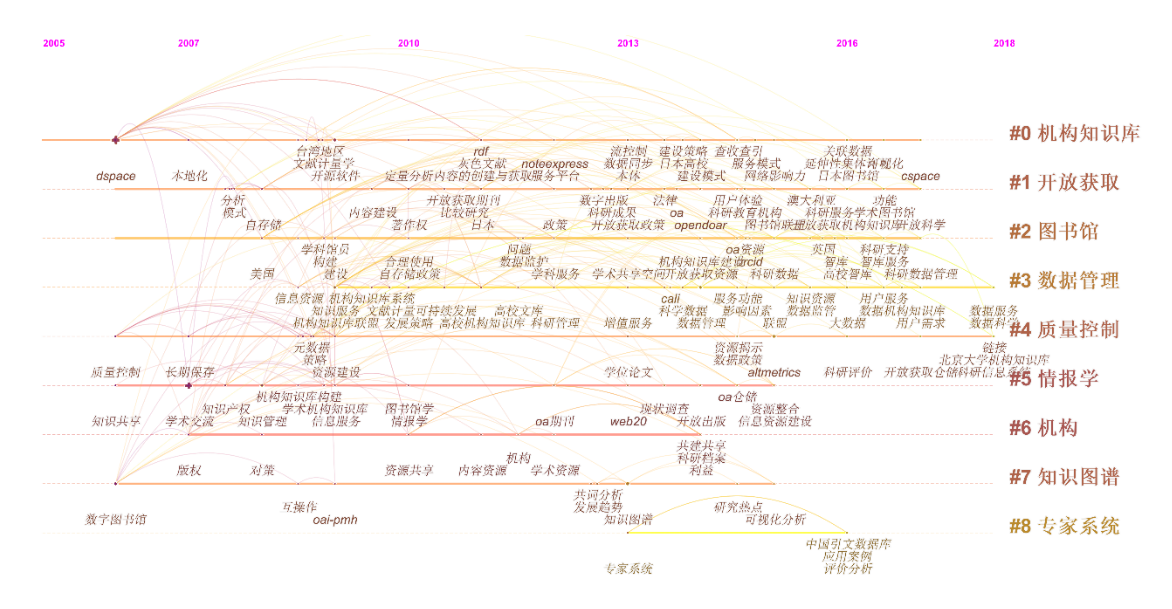


图 6 关键词时间线视图

纵观 9 个类别的时间轴线，只有 #3 和 #4 的时间轴线（实线部分）延续至 2018 年，说明这两类属于当前的研究热点。结合其他时间轴上新出现的关键词，可以概括出国内 IR 研究新出现的研究倾向，这些倾向可能预示着新的研究热点，主要包括如下几个方面：

（1）科研数据。新出现的关键词中“数据”出现的频率最高，说明国内 IR 研究已经开始关注科研数据领域。科研数据属于 IR 资源建设范畴，在过去的有关 IR 资源建设研究中，主要关注点是期刊论文、会议论文、演示文稿、学位论文点等比较有“成果”性质的显性学术资源。科研数据在科研活动中往往只是“中间产品”，或者只是学术成果的附属资料，常常被科研工作者忽略，随着科研活动的结束而被放弃。然而随着科研活动进入第四范式，以及计算科学领域的进步，特别是在大数据研究背景下，数据重用已经被认可和重视，促使人们开始注重科研活动中有价值数据的保存与重用 [16]。

（2）科研服务。IR 因科研而生，并终将为科研服务。科研活动是动态的、持续的过程，并且随着开放科学的进步和社会经济环境等因素的变化而不断变化和调整。同样地，为科研服务的 IR 也将不断自我调整和完善，也是今后 IR 研究中的一个持续热点问题。

（3）辅助科研决策。随着大数据科学的发展，以及 IR 在内容和功能上的不断完善，IR 在机构科研决策方面的辅助决策权重上的占比将越来越高，将扮演重要的“智库”角色。当然，辅助科研决策并不单单基于单个机构的知识库，而可能在多个机构的“IR 联盟”之上。IR 参与辅助决策是国内 IR 研究和建设者一个不断追求的目标，需要长期的不断努力与争取。

（4）数据增值服务。IR 经过多年发展，在内容上几乎囊括了本机构的所有公开发表的学术成果以及部分未公开发表的学术成果，并开始涉及科研数据。因此，IR 将在学术数据上占有优势，除了提供常规的基础服务，还将在高校“双一流”的建设热潮中，发挥数据优势，提供数据增值服务 [17]。

另外，CSpace 作为中科院图书馆兰州分馆推出的中国本土化的 IR 构建软件系统，在中科院系统内得到了很好的推广和使用。随着软件与时俱进的更新与完善，并不断拓展用户群体，今后有关 CSpace 的发文可能会保持一定的频率。

6 分析与探讨

通过前文对 1 362 篇论文从不同角度可视化分析后，可以基本了解我国目前 IR 研究的基本态势，即我国的 IR 研究与建设已经取得了不俗的成就，现在正进入一个比较关键的关口。笔者针对国内 IR 研究与建设的当前现状，提出 4 点建议为 IR 今后发展提供参考。

6.1 再接再厉，持续推进机构知识库建设

据统计[17]，大陆地区高校和科研院所已建与在建 IR 共有 325 个，资源类型包括期刊论文、学位论文、会议论文、图书、专利、研究报告、演示课件、实验数据等。换句话说，目前我国很多高校和科研院所正在建设一个规模比较庞大的学术成果资源库，是学术“大数据”，随着科技的发展，国家对数据越来越重视，已经把大数据提升为国家基础战略资源，并认为大数据是信息社会的“石油”[18]，在此背景下，继续推进 IR 建设具有重大“战略”意义，发展和应用前景广阔。

另一方面，IR 作为持续收集、管理机构教学科研活动中产生的智力成果平台，只要教学科研活动不停止，IR 建设也不应该停歇，应该从政策、资金等多方面持续推进 IR 建设。

6.2 整合优化，定位成为“国家基础战略资源”的组成部分

目前国内 IR 建设成绩斐然，构建了一大批以机构学术成果为主要内容的 IR 系统，资源总量也达到相当规模，并出现了以资源共享为主要目的的两大 IR 资源联盟（“中国科学院文献情报中心知识库”和“中国高校机构知识库联盟”）。应该顺应国家发展形势，将这些数量规模可观并且还在不断增长的 IR 资源，定位为“国家基础战略资源”的组成部分，从国家的层面和高度来推动 IR 资源的建设与开发。然而，目前这些资源还不能实现真正的互联和共享，仍分散在不同教育机构和科研院所，资源主体还是掌握在各个实体机构手中。

资源只有在流动中才能产生其价值，因此需要对当前 IR 资源进行整合与优化，使其互联共享。然而，对全国性的 IR 资源进行彻底的整合优化，必然涉及不同教育科研机构之间的各种机制冲突和利益调整，而且这个过程不可能由某一个机构来完成，而是需要从国家层面，对现有的 IR 政策、体制、机制等方面进行改革创新，对多方利益进行协调与平衡。而 IR 建设者需要从大局出发，解放“各自为战”思想，拓宽目前对 IR 资源开发与利用的思路和视域。只有对 IR 资源进行比较彻底的整合与优化，才能成为真正的“国家基础战略资源”。

6.3 统筹协作，促进机构知识库向纵深发展

从前文的热点研究分析中可知，相当多的来自不同单位的作者进行着相同或相似的 IR 主题研究，两大 IR 研究群体（中科院群体和高校群体）也在进行相类似的 IR 研究和 IR 联盟建设。由于 IR 研究人员或团体相对分散，缺少统一规划和统筹，导致 IR 研究热点广泛，但又不能形成真正的“热点”，同时造成重复投入、重复研究、重复产出、重复劳动、重复构建等现象，浪费了人员、精力，甚至财力。这些问题在 IR 早期自发的、粗犷式、“无政府主义”发展模式中，难以避免。

IR 发展到今天，必须对此前的发展模式进行改革，对全国 IR 研究团体和个人及其研究项目进行全面统筹、分工协作，避免重复研究。各 IR 研究团体和个人应结合各自特点和优势有针对性地对 IR 发展趋势中的热点问题展开研究，例如科研数据热点研究、科研辅助热点研究等，形成有一定特色的 IR 热点研究。同时，从国家层面，组织一批有实力的 IR 研究团队对 IR 发展趋势中的热点问题进行预研攻关性研究，例如，IR 资源作为“国家基础战略资源”，探索

知识管理论坛，2019(5):276-286

DOI：10.13266/j.issn.2095-5472.2019.027

如何利用大数据的技术方法对其进行深度开发与利用，从而使我国 IR 研究具有一定特色，又形成一定梯度的 IR 研究队伍。

通过全面统筹和规划协调，能够节约、经济、高效地实现 IR 研究与建设的突破与创新，促使国内 IR 向纵深发展。

6.4 以评促建，提升机构知识库构建质量和水平

如前所述，国内早期 IR 建设较为粗犷，缺少指导和统一规划，IR 资源种类繁多以及机构对 IR 建设投入多寡不一，致使 IR 普遍存在系统功能较单一、资源种类不一、资源著录不规范、不详细等问题。这些问题会对后续的 IR 系统升级和资源建设与利用造成影响。

为了使 IR 更好更快地发展，使 IR 资源更好地互联共享，建立 IR 建设质量评价体系是一种行之有效的手段。从文献调研来看，国内外已经有学者提出有关 IR 质量评价的指标体系，R. Serrano-Vicente 等提出从技术、规程、内容、营销和人员角度对西班牙的 IR 进行评价 [19]，郭清蓉提出并建立 IR 评价体系的思路和具体指标 [20]，郭翊提出由 9 项原则及其所包含的绩效指标构成的 IR 评价体系 [21]，唐奇提出基于平衡计分卡的 IR 评价指标体系 [22]。

然而，IR 建设是一个动态过程，对其进行评价也需要灵活的动态的评价体系。因此，需要从国家战略目标高度，对 IR 建设的不同阶段采用不同评价指标体系进行及时评估，确保 IR 向高质量、高水平建设迈进。

7 结语

由于本研究数据来源仅限于 CNKI 收录的期刊论文，检索用词是“机构知识库”及其相近概念词语，并且未包括博硕士学位论文在内的其他学术研究数据，研究结果可能会与实际情况有所偏差。另外，其他具有一定计量分析意义的元数据字段（如基金资助、作者单位、参考文献等），由于来源数据不规范在本研究中未被纳入，希望在后续研究中将其纳入并分析。

笔者对 IR 未来发展的建议是立足于宏观的国家层面而提出，期望国家相关部门出面引领统筹国内 IR 未来发展，组织国内 IR 研究专家、IR 核心刊物编辑和审稿专家，以及相关大数据应用专家等人员，成立 IR 顾问决策小组，在总结梳理过去 IR 建设经验基础上，对 IR 未来发展作顶层规划和设计，确定发展战略，配套出台相关政策等。前文所述的 IR 发展建议可能带有一定的理想性，仅借此抛砖引玉，希望引起国内 IR 研究同仁对 IR 未来发展作更多讨论与探索。

参考文献：

[1] 钟源 . 2004-2013 年国外机构知识库研究的文献计量分析 [J]. 情报科学 , 2016, 34(6):71-76.

[2] DURASPACE. DSpace - A Turnkey Institutional Repository Application [EB/OL].[2019-03-30]. https://duraspace.org/dspace/.

[3] MIT Libraries. DSpace@MIT[EB/OL].[2019-03-30]. http://dspace.mit.edu/.

[4] 厦门大学图书馆 . 厦门大学学术典藏库 [EB/OL].[2019-03-30]. https://dspace.xmu.edu.cn/.

[5] 祝忠明 , 马建霞 , 常宁 , 等 . 基于 DSpace 构建学科知识库系统的研究与实践 [J]. 现代图书情报技术 , 2006(7):10-14.

[6] 上海市图书馆学会 . 历届数字图书馆前沿问题高级研讨班 [EB/OL].[2019-03-30].http://society.library.sh.cn/node/2827.

[7] 中国机构知识库推进工作组 . 历届会议 [EB/OL].[2019-03-30].http://2018chinair.csp.escience.cn/dct/page/70002.

[8] ROY T. Institutional repositories[J]. Library journal, 2002, 127(15): 28-30.

[9] 吴建中 . 图书馆 VS 机构库——图书馆战略发展的再思考 [J]. 中国图书馆学报 , 2004(5):7-10.

[10] 祝忠明 , 张旺强 , 刘巍 , 等 . 我国数字图书馆软件开源实践的策略和模式：以 CSpace 为例 [EB/OL].[2019-03-30]. http://dlib-oss2014.csp.escience.cn/dct/attach/Y2xiOmNsYjpwZGY6MjU1NzY=.

[11] 聂华 , 韦成府 , 崔海媛 .CALIS 机构知识库 : 建设与推广、反思与展望 [J]. 中国图书馆学报 , 2013, 39(2):46-52.

[12] 韦成府 , 聂华 , 崔海媛 . 多馆协作开发的机构知识库建设——以 CALIS 机构知识库项目为例 [J]. 大学图书馆学报 , 2014, 32(3):69-73.

[13] 李国俊 , 王瑜 , 王李梅 , 等 . 基于元数据的高校机构知



识库建设研究——以北京科技大学机构知识库为例 [J]. 大学图书馆学报 , 2012, 30(4):55-60.

[14] 李春旺 . 网络环境下学术信息的开放存取 [J]. 中国图书馆学报 , 2005(1):33-37.

[15] 卫垌圻 , 范广兵 , 翁硕 , 等 . 研究型图书馆机构知识库资源的利用与增值服务 [EB/OL]. [2019-03-30]. http://ir.psych.ac.cn/handle/311026/27283.

[16] 陈和 . 刍议国内高校科学数据知识库的构建 [J]. 数字图书馆论坛 , 2017(12):45-51.

[17] 龚亦农 , 朱茗 . 我国机构知识库建设现状调查 [J]. 数字图书馆论坛 , 2018(9):20-28.

[18] 杨波 . 大数据成国家基础战略资源 政府数据互联共享促产业腾飞 [EB/OL].[2019-03-30].http://www.gov.cn/zhengce/2015-08/20/content_2916507.htm.

[19] SERRANO-VICENTE R, MELERO R, ABADAL E. Evaluation of Spanish institutional repositories based on criteria related to technology, procedures, content, marketing and personnel[J]. Data technologies and applications, 2018, 52(3): 384-404.

[20] 郭清蓉 . 机构知识库自存储评价体系的构建 [J]. 情报杂志 , 2009, 28(7):74-76, 73.

[21] 郭翊 , 吉萍 . 机构知识库评价体系的构建 [J]. 图书馆学刊 , 2014, 36(2):10-12.

[22] 唐奇 . 基于平衡计分卡的机构知识库评价指标体系构建 [J]. 图书馆工作与研究 , 2019(4):53-58.

作者贡献说明：

陈 和：提出研究思路，收集与处理数据，撰写、修改论文和定稿；

林 静：处理数据与制图，撰写论文。

Visual Analysis of Domestic Institutional Repository in 2000-2018

Chen He Lin Jing

Library of Xiamen University, Xiamen 361005

Abstract: [Purpose/significance] At present, IR has reached a relatively mature stage. This article combs, analyzes and summaries the development of domestic IR, to provide reference for the next stage of IR development. **[Method/process]** Using the software of Excel, VOSviewer, CiteSpace and programming, 1 362 articles of IR were collected from CNKI in 2000-2018 year for visual analysis. **[Result/conclusion]** The results of this study are as follows: IR articles publishing had experienced rapid development period, the relative stationary period, and downturn period, which means that the current research of domestic IR has reached a new stage, and have new research hotspots; the high-frequency authors (groups) are mainly composed of the authors of the Chinese Academy of Sciences and universities, the former's research is more advanced and deeper than the latter; "Library and Information Service", "Journal of Modern Information", "Researches in Library Science", etc. are the most published journals with IR research papers, and are more influential in the field of IR research in China; research hotspots of domestic IR include system construction, resource construction, academic communication, knowledge management and services, copyright protection, etc.; future research hotspots include research data, scientific research services, auxiliary research decisions, and data value-added services.

Keywords: institutional repository IR data visualization mapping knowledge domain bibliometric research hotspots